

Measuring inflammatory foot disease in rheumatoid arthritis: development and validation of the Rheumatoid Arthritis Foot Disease Activity Index-5

Hoque, Anika; Gallagher, Kellie; McEntegart, Anne; Porter, Duncan; Steultjens, Martijn; Woodburn, James; Hendry, Gordon J.

Published in:
Arthritis Care & Research

DOI:
[10.1002/acr.24259](https://doi.org/10.1002/acr.24259)

Publication date:
2021

Document Version
Author accepted manuscript

[Link to publication in ResearchOnline](#)

Citation for published version (Harvard):

Hoque, A, Gallagher, K, McEntegart, A, Porter, D, Steultjens, M, Woodburn, J & Hendry, GJ 2021, 'Measuring inflammatory foot disease in rheumatoid arthritis: development and validation of the Rheumatoid Arthritis Foot Disease Activity Index-5', *Arthritis Care & Research*, vol. 73, no. 9, pp. 1290-1299.
<https://doi.org/10.1002/acr.24259>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please view our takedown policy at <https://edshare.gcu.ac.uk/id/eprint/5179> for details of how to contact us.

DR. GORDON JAMES HENDRY (Orcid ID : 0000-0003-1611-7193)

Article type : Original Article

Title: Measuring inflammatory foot disease in rheumatoid arthritis: development and validation of the Rheumatoid Arthritis Foot Disease Activity Index-5

Authors: Anika Hoque (BSc Hons)^{1 2}, Kellie Gallagher (PhD)³, Anne McEntegart (MBCChB)⁴, Duncan Porter (MD)⁵, Martijn Steultjens (PhD)¹, James Woodburn (PhD)¹, Gordon J Hendry^{1*}(PhD)

Affiliations: ¹School of Health & Life Sciences, Glasgow Caledonian University, Cowcaddens Road, Glasgow, G4 0BA, UK; ²Podiatry East Quadrant, NHS Greater Glasgow and Clyde, Trust Headquarters, Balgrayhill Road, Glasgow, G21 3UR, UK; ³Health, Sport & Bioscience, University of East London, Water Lane, E15 4LZ; ⁴Stobhill Hospital, 133 Balornock Rd, Glasgow G21 3UW; ⁵Gartnavel General Hospital, 1053 Great Western Rd, Glasgow G12 0YN.

***Corresponding author contact details:**

Dr Gordon J Hendry

Phone: +44 141 331 3635

Email: gordon.hendry@gcu.ac.uk

Word count: 3,793

Financial support: Anika Hoque was supported via the Versus Arthritis Musculoskeletal Allied Health Professionals Internship Scheme.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/acr.24259](#)

This article is protected by copyright. All rights reserved

Objective: Omission of foot joints from composite global disease activity indices may lead to underestimation of foot and overall disease in rheumatoid arthritis (RA) and under-treatment. The aim of this study was to evaluate the measurement properties of the Rheumatoid Arthritis Foot Disease Activity Index-5 (RADAI-F5), a newly developed patient-reported outcome measure for capturing foot disease activity in people with RA.

Methods: Participants with RA self-completed the RADAI-F5, modified Rheumatoid Arthritis Disease Activity Index (mRADAI-5), Foot Function Index (FFI) and Foot Impact Scale impairment/footwear and activity/participation subscales (FIS-IF and FIS-AP). The 28 joint Disease Activity Score (DAS28-ESR) was also recorded. Subgroups completed the RADAI-F5 at 1 week and 6-months. Psychometric properties including construct, content, and longitudinal validity, internal consistency, 1-week reproducibility, and responsiveness over 6-months were evaluated.

Results: Of 142 respondents, 103 were female, with a mean (standard deviation [SD]) age of 55 years (12.5) and median (interquartile range [IQR]) RA disease duration of 10 (3.6-20.8) months. Theoretically consistent associations confirming construct validity were observed with mRADAI-5 (0.789, 95% confidence interval [CI] 0.73-0.85), FFI (0.713, CI 0.62-0.79) and FIS-IF (0.695, CI 0.66-0.82) ($p < 0.001$); FFI-AP (0.478, $p < 0.001$, CI 0.37-0.63) and the DAS28-ESR (0.379, $p < 0.001$, CI 0.26-0.57). The RADAI-F5 demonstrated high internal consistency (Cronbach's Alpha=0.90) and good reproducibility (ICC=0.868, $p < 0.001$, CI 0.80-0.91; smallest detectable change=2.69). Content validity was confirmed with 82% rating the instrument relevant and easy to understand.

Conclusion: The Rheumatoid Arthritis Foot Disease Activity Index-5 is a valid, reliable, responsive, clinically feasible PROM for measuring foot disease activity in RA.

Significance and Innovations

- The RADAI-F5 is the first patient-reported outcome measure designed to measure localised foot disease activity in rheumatoid arthritis.

- The RADAI-F5 is valid, reliable, responsive to change, and is feasible for use in clinical practice to measure foot disease in people with rheumatoid arthritis.
- The RADAI-F5 provides a means for measuring foot disease activity which is not captured by composite global disease activity indices such as the DAS28-ESR.

Foot pain and disability are common in people who have rheumatoid arthritis (RA) with up to 90% experiencing disease-related foot symptoms during the course of their disease [1-4]. Synovitis in the small joints of the feet is present at the onset of RA in up to 70% of patients [5]. With the implementation of pharmacological management, prevalence of forefoot disease stabilises at 40-50% after 2 years, but the prevalence of radiographic joint damage increases from 20% to 50% [5]. Approximately half of patients experience hindfoot joint problems [6], and clinically important soft tissue disease such as tibialis posterior tenosynovitis has a reported prevalence ranging from 13-64% [3].

The management of RA involves a 'treat to target' strategy that comprises regular review, objective assessment of disease activity and escalation of treatment if there is persistent disease activity. The aim is to achieve clinical remission (or low disease activity) [7,8]. Evaluation of inflammatory disease activity in RA involves the composite disease activity indices, of which the most widely used are the 28 joint count disease activity score (DAS28-ESR) [9], the clinical disease activity index (CDAI), and the simplified disease activity index (SDAI) [10]. These include 28-joint counts for tenderness and swelling that omit the joints and soft tissues of the feet [11]. Recent studies have demonstrated that approximately one-third of patients with RA classed as in DAS-28, SDAI and/or CDAI remission had clinically determined active foot synovitis [2,11,12]. Although the examination of foot joints is recommended [13], patients treated solely according to disease activity indices may be at risk of ongoing foot joint damage [11,14].

Various self-reported measures of RA disease activity have been validated including two versions of the Rheumatoid Arthritis Disease Activity Index (RADAI and modified RADAI-5) [15,16]. While these tools are more likely to capture foot disease activity than composite indices that exclude foot joints examination, they are not widely used in clinical practice. Several questionnaires have been developed for measuring RA-related foot problems including the Foot Impact Scale (FIS), Foot Function Index (FFI), and the Salford Rheumatoid Arthritis Foot Evaluation instrument (SAFE) [17-19]. These largely focus on disability as opposed to disease activity, and so have limited value for informing medical management. In addition, these tools lack feasibility in clinical practice due to high number of items and associated time burden, as well as absence of clinically meaningful categories for ease of interpretation.

The current gold standard technique for measuring foot disease activity is magnetic resonance imaging (MRI) [20]. The more widely used imaging alternative is musculoskeletal ultrasound (MSUS), which is less sensitive than MRI and more specific than clinical examination of disease activity in RA [20-22]. However, these methods are not routinely used due to impracticalities such as cost of scans/equipment, time to perform scans, training needs and the risk of exposure to contrast agents. There is also a lack of consensus concerning how to approach medical management of subclinical disease detected using MSUS [23].

In summary, there is no widely used, validated and clinically feasible method for the assessment of foot disease activity in RA. We propose that a PROM designed to measure local foot disease activity may provide an opportunity for a treat-to-target medical approach which does not exclude foot disease activity. Accordingly, the purpose of this study was to develop and validate a new concise measure of foot disease activity for people with RA.

PATIENTS AND METHODS

Development of the questionnaire

The design and content of the RADAI-F5 was derived from the modified Rheumatoid Arthritis Disease Activity Index-5 (mRADAI-5), a 5-item patient reported outcome measure for the self-report of global disease activity, developed and evaluated by Leeb et al [16] and Rintelen et al [24]. It is completed in a numerical rating scale (NRS) format from 0 to 10, and scored by an average summary score ranging from 0-10. The RADAI-F5 was developed by editing the mRADAI-5 with an opening statement: "THINKING ONLY OF YOUR FEET", and editing the original questions to subsequently read as follows: "How active was your arthritis IN YOUR FEET over the last 6 months?" (0 = completely inactive to 10 = extremely active); "How active is your FOOT arthritis today with respect to joint tenderness and swelling?" (0 = completely inactive to 10 = extremely active); "How severe is your arthritis pain IN YOUR FEET today?" (0 = no pain to 10 = unbearable pain); "How would you describe your general FOOT health today?" (0 = very good to 10 = very bad); "Did you experience foot joint stiffness on awakening yesterday morning? If yes, how long was this stiffness IN YOUR FEET?" (0 = no stiffness to 10 = stiffness the whole day). The RADAI-F5 is scored by an average summary score ranging from 0-10.

Study setting and participants

The two data sources for this study were 1) a primary RADAI-F5 validation study, conducted at rheumatology outpatient clinics at Glasgow Royal Infirmary, Gartnavel General Hospital and Stobhill Hospital within the Greater Glasgow and Clyde National Health Service Board, and 2) a larger randomised controlled trial, the details of which have been published previously [25]. Briefly, the trial was a multi-centre, parallel-group, randomised control trial with a 6- and 12-months follow-up periods, with participants randomly allocated to either customised or prefabricated foot orthoses. Trial participants were recruited from rheumatology outpatient clinics

within NHS Grampian, Fife, Lanarkshire, Lothian Health Boards, Dorset Healthcare University Trust, and Homerton University Hospital Trust.

Participants were included if they were aged between 18 and 75 years of age, with a definitive clinical diagnosis of RA. Patients were excluded if they were unable to read, write and/or understand the English language, or if they were diagnosed with other major medical conditions which could have diminished ability to distinguish between RA-related foot problems and those due to alternative disease mechanisms. Ethical approval was obtained from the West of Scotland Research Ethics Committee 5 (13/WS/0106) and the East of England Essex Research Ethics Committee (15/EE/0410). Participants were recruited consecutively and written consent was obtained from all participants.

Data collection and measures

Demographic and clinical information was collected as baseline including age, sex and disease duration. The newly developed RADAI-F5 was collected at baseline, 1-week from baseline, and 6-months from baseline. All other measurements were recorded at baseline and 6-months. The DAS28-ESR scores were recorded by rheumatologists as part of routine care and made available to researchers. The mRADAI-5 was collected as an additional self-reported measure of global disease activity [16]. Foot related impairments and disability were evaluated using the Foot Function Index (FFI) [18], and the Foot Impact Scale [17]. The FFI is a widely used and extensively validated 23-item PROM, completed using a 100mm VAS format, providing a mean summary score from 0-100 (higher scores indicating worse disability) [18]. The FIS is an extensively validated RA-specific 51-item measure with domains for impairment/footwear (21-items) and activity limitation/participation restriction (30-items). It is completed using a yes/no dichotomous format and scores for domains calculated by summing “yes” responses (higher scores indicating worse disability) [17].

In order to evaluate the content validity and practical burden of the RADAI-F5 an additional 3 items were evaluated; a 5 point Likert scale regarding questionnaire relevance to participants (ranging from extremely irrelevant to extremely relevant), a 5 point Likert scale regarding participants' opinion on the readability/understanding of the new questionnaire (ranging from very difficult to very easy), and time taken to complete the questionnaire (in minutes).

Analysis

Data were analysed using SPSS 25 (IBM) and Excel 2016 (Microsoft). Descriptive statistics for age (median, inter-quartile range) in years, sex (female to male ratio) and disease duration (median, inter-quartile range) in months were generated for all participants at baseline.

The RADAI-F5 was examined using factor analysis by principal component analysis to reveal the structure and item loading. Kaiser-Meyer-Olkin test and Bartlett's test of sphericity were undertaken to determine data suitability for factor analysis. The number of factors extracted was decided by a combination of Kaiser's rule (eigenvalues greater than 1), examination of the scree plot, and interpretation of items' contribution to the factor.

To test internal consistency, we evaluated the inter-item correlation matrix and calculated Cronbach's alpha, a measure of consistency between items in a scale. A Cronbach's alpha of 0.7-0.9 was considered acceptable [26,27].

Hypotheses were generated *a priori* to examine the extent to which baseline scores (construct validity) and 0-6 month change scores (longitudinal validity) on the RADAI-F5 were associated with baseline and 0-6 month change scores from other measures in a manner that was theoretically consistent [28]. Hypotheses for construct validity, which focused on baseline scores, were specified as follows: moderate positive correlations between the RADAI-F5 score and mRADAI-5, FFI, and FIS domains; and a positive weak correlation between the RADAI-F5 score and the DAS-28 score. Hypotheses for longitudinal validity, which focused on 0-6 month change scores, were identical except for the FIS subscales, where a weak positive correlation was anticipated as the FIS is less responsive to change [29]. Spearman's rank (r_s) correlation and 95% confidence intervals (95%CI) were used to test these hypotheses and coefficients were interpreted as follows: 0-0.1 = negligible, 0.1-0.39 = weak, 0.4-0.69 = moderate, 0.7-0.89 = strong, 0.9-1.0 = very strong [30].

The 1-week (test-retest) reliability was examined using a two-way mixed intra-class correlation coefficient (ICC) with corresponding 95% CIs for baseline and 1-week scores. Once preliminary foot disease categories were established (see below), Cohen's quadratic weighted kappa and

corresponding 95%CI for foot disease categories (remission, low, moderate, high) was calculated, with values of greater than 0.61 indicating substantial reliability [31].

Absolute measurement error was evaluated using the standard error of measurement (SEm), derived by dividing the SD of the mean change between the two measurements ($SD_{\text{change}}/\sqrt{2}$); the 95% limits of agreement (LOA), derived by calculating the mean change between the two measurements, plus or minus 1.96 times the standard deviation of the changes ($(\text{mean}_{\text{change}}) \pm 1.96 \times (SD_{\text{change}})$); the 95% smallest detectable change (SDC_{95}) ($1.96 \times \sqrt{2} \times \text{SEm}$), and construction and examination of Bland-Altman plots [32-34].

Responsiveness was evaluated using four different effect size statistics: Wilcoxon signed ranks test, Cohen's d, the standardised response mean (SRM) and Guyatt's Index [35]. In the absence of an anchor question to calculate minimal important difference (MID), MID was calculated using a value of $0.5 \times SD_{\text{change}}$ scores between baseline and 6-months. Guyatt's Index, representing the magnitude and variability in change scores relative to its MID was calculated as $MID/\sqrt{(2 \times SD_{\text{change}})}$ [36]. Effect sizes were interpreted as follows: <0.15 = negligible, >0.15 - <0.40 = small, ≥ 0.40 - <0.75 = medium, ≥ 0.76 - <1.10 = large, ≥ 1.10 - <1.45 = very large, and >1.45 = huge [35].

Participants were classified according to mRADAI-5 thresholds for remission, mild, moderate or high disease activity. With participants assigned to the mRADAI-5 reference categories, the third quartile of corresponding RADAI-F5 scores were calculated to establish the thresholds for respective RADAI-F5 categories [24]. Cohen's quadratic weighted kappa and 95% CI was used to evaluate agreement between disease activity categories between the mRADAI-5 and the RADAI-F5.

Median (IQR) values were obtained for readability and relevant Likert scores, and completion time. For evaluation of floor and ceiling effects for the RADAI-F5 in the RA population, we adopted the conventional 15% threshold for patients achieving the highest and lowest scores to define and ceiling and floor effect respectively [32,37].

For the purposes of evaluating structural validity via factor analysis, a minimum sample size of $n \geq 100$ at baseline was targeted *a priori* in order to achieve a participant to item ratio of 20 [38].

For hypotheses testing for construct and longitudinal validity, between 61 and 123 participants were required to detect at least weak correlation coefficients from 0.25-0.35 at 80% power and 0.05 significance level (G*Power 3.1.9.2).

RESULTS

A total of 142 participants (72.5% female) with a median (IQR) age of 55.5 (50-62) years and median (IQR) disease duration 10 (3.6-20.8) months took part, including 37 from the primary RADAI-F5 study and 105 from the RCT. A total of 84 participants completed the RADAI-F5 at 1-week for reproducibility analyses, and 64 completed 6-month follow-ups for responsiveness analyses. Median (IQR) DAS28-ESR and mRADAI-5 scores indicate that participants were typically in a moderate disease activity state at study baseline (Table 1). For DAS-28-ESR disease categories, 21.5% were in remission, 8.4% had low disease activity, 47.7 had moderate disease activity and 22.4% had high disease activity. Median (IQR) FFI, FIS impairment/footwear and activity limitation/participation restriction subscale scores suggest participants typically presented with moderate foot-related disability at baseline (Table 1).

Table 1. Sample baseline characteristics.

Dimensionality

The Kaiser-Meier-Olkin value of 0.83 and highly significant Bartlett test ($p < 0.001$) indicate sampling adequacy and suitability for structure detection and factor analysis. Both Kaiser's 'eigenvalue greater than one' rule and scree plot examination suggested a one-factor solution, and this explained 73.18% of the common variance. Item loadings on the factor were uniformly greater than 0.4, indicating that all items contributed significantly to the aggregate score.

Internal consistency

High inter-item correlations (>0.8) were observed for questions 2 and 3, 2 and 4 and 3 and 4 of the questionnaire. Moderate inter-item correlations (>0.6) were observed for question 1 and 2, and 2 and 5. All other inter-item correlations were between 0.4-0.6. Cronbach's alpha was 0.90, indicating a high level of internal consistency.

Construct validity

The RADAI-F5 had a weak positive correlation with the DAS-28 (r_s 0.38, 95% CI 0.26-0.57, $p<0.001$) (Figure 1); a moderate positive correlation with the FIS impairment/footwear (r_s 0.69, 95% CI 0.66-0.82, $p<0.001$) and FIS activity limitation/participation restriction subscales (r_s 0.48, 95% CI 0.37-0.63, $p<0.001$). Stronger positive correlations than predicted were observed for the mRADAI-5 (r_s 0.79, 95% CI 0.73-0.85, $p<0.001$), and the FFI (r_s 0.71, 95% CI 0.62-0.79, $p<0.001$) (Figure 1). Sixty per cent of associations for construct validity were in line with *a priori* hypotheses, and therefore largely theoretically consistent. Construct and longitudinal validity analyses are presented in Table 2.

Figure 1.

Longitudinal validity

The RADAI-F5 had a weak positive correlation with the DAS28-ESR (r_s 0.33, 95% CI 0.04-0.52, $p=0.011$); a moderate positive correlation with the mRADAI-5 (r_s 0.66, 95% CI 0.41-0.74, $p<0.01$), the FFI (r_s 0.60, 95% CI 0.47-0.77, $p<0.001$) and the FIS impairment/footwear subscale (r_s 0.43, 95% CI 0.09-0.56, $p=0.001$). A weaker positive correlation than predicted was observed for the FIS activity limitation/participation restriction subscale (r_s 0.19, 95% CI -0.01-0.49, $p=0.156$), and the FFI (r_s 0.71, 95% CI 0.62-0.79, $p<0.01$). Eighty per cent of associations for longitudinal validity were in line with *a priori* hypotheses.

Table 2. RADAI-F5 construct and longitudinal validity correlations for *a priori* hypotheses testing for associations with alternative disease activity and foot disability measures.

Reliability

Mean (SD) RADAI-F5 scores for baseline and 1-week were 4.8 (2.58) and 4.91 (2.74). The mean (SD) difference between baseline and 1-week for the RADAI-F5 was 0.11 (1.37). The ICC (95%) for 1-week reproducibility was 0.87 (0.80-0.91, $p<0.001$), indicating very good reproducibility.

Absolute measurement error

The SEM value calculated for the RADAI-F5 from baseline and 1-week data was 0.97. This can be interpreted as follows; if a patient scores 3 on the RADAI-F5, we can be 68% confident that their true score lies between 2.03 and 3.97, and 95% confident that their true score lies between 1.1 and

4.9. The LOA were -2.57 and 2.80. From the Bland-Altman plot (Figure 2), we can identify that 97.6% of differences between the two time-points were within the 95% limits. Exploration of the plot suggests there may be a minor funnel effect, with spread increasing slightly with increasing mean concentration (higher foot disease activity). The SDC95 value was estimated as 2.69, representing 26.9% of the RADAI-F5 scale maximum range, meaning that we can be 95% confident that a change score of 2.69 or more is a true change.

Figure 2.

Responsiveness

The RADAI-F5 exhibited high responsiveness to change over 6 months, as evidenced by a highly significant Wilcoxon signed-ranks tests ($p < 0.001$) with very large effect size (0.91), Cohen's $d = 0.64$ (medium effect size), and an SRM = 0.97 (very large effect size). The MID value obtained from the distribution method ($0.5 \times SD_{\text{change}}$) was 1.16, and the subsequent GI value calculated was 0.70 (medium effect size). These values can be interpreted as a consistent pattern of medium to high responsiveness for the RADAI-F5.

Interpretability

For remission according to the mRADAI-5, the RADAI-F5 median and third quartile were 0.73 and 1.0 ($n=11$). Therefore, the remission state was defined as a RADAI-F5 score ranging from 0-1.0. The same procedure applied to define the disease categories for mild, moderate and high disease activity, resulted in the following ranges: $>1-3.6$ for mild disease activity ($n=15$), $>3.6-5.7$ for moderate disease activity ($n=42$), and $>5.7-10$ for high disease activity ($n=65$) (table 3). Agreement between the RADAI-F5 and mRADAI-5 categories was good, as expressed by 72.5% exact agreement and Cohen's quadratic weighted kappa (95% CI) of 0.71 (0.56-0.85). For the newly derived RADAI-F5 disease categories, agreement between baseline and 1-week was good, as expressed by 70.0% exact agreement and Cohen's quadratic weighted kappa (95% CI) of 0.81 (0.75-0.88). Characteristics of disease and foot-related disability status within newly developed RADAI-F5 disease categories are shown in table 3.

Table 3. Participant characteristics according to RADAI-F5 foot disease activity categories (median and IQR unless specified).

Content validity and practical burden

Participants largely considered the RADAI-F5 to be relevant for measuring foot disease activity in RA, with 53% and 29% indicating that it was relevant and extremely relevant respectively (Figure 3A). Participants largely considered the RADAI-F5 to be easy to read and understand, with 45.8 and 40% indicating that it was easy and very easy to read/understand respectively (Figure 3B). The median (IQR) time to complete the questionnaire was 5 (2-15) minutes. Overall it appears that the RADAI-F5 has good content validity, low completion burden.

Figure 3.

Floor/ceiling effects

A total of 6 (4.33%) participants achieved the lowest possible RADAI-F5 score (0), and 0 (0%) participants achieved the highest possible score (10). Based on the 15% threshold levels, the RADAI-F5 does not exhibit a ceiling or floor effect.

DISCUSSION

We have developed and validated a 5-item PROM, named the Rheumatoid Arthritis Foot Disease Activity Index (RADAI-F5) to allow for the monitoring of inflammatory foot disease activity in people with established and early RA. The psychometric properties meet the recommended standards set by the International Society for Quality of Life Research (ISOQOL) [39] and the Consensus Based Standards for the Selection of Health Measurement Instruments (COSMIN) [40] demonstrating good construct validity, reliability, content validity, internal consistency, responsiveness and interpretability. The new PROM is designed to be quick and simple for patients to complete and clinicians to score and interpret, so that it can be used alongside composite disease activity indices in clinical practice. We anticipate that future use of the RADAI-F5 as an adjunct to composite disease activity indices will improve local disease monitoring and may facilitate better medical management of foot disease activity. The RADAI-F5 may also be used alongside existing disease-specific foot disability PROMS such as the FIS in rehabilitation settings to distinguish between inflammatory and mechanical/functional foot impairments, which may help to inform new and extended paradigms of foot care in early RA [41]. Moreover, it is

possible that the RADAI-F5 could negate the need for composite disease activity scores to be recorded for RA foot research purposes (often collected to account for confounding).

A key finding of this study is that evidence of convergent validity was observed between the RADAI-F5 and the mRADAI-5, and divergent validity between the RADAI-F5 and the DAS28-ESR. This finding was anticipated *a priori* and is theoretically consistent with existing evidence demonstrating that composite disease activity indices that omit foot joints do not adequately capture foot synovitis [11]. Indeed, the majority of our *a priori* specified hypotheses were confirmed for strength of associations between disease activity/foot disability measures with the RADAI-F5. The correlation between RADAI-F5 and mRADAI-5 was slightly stronger than anticipated, but perhaps unsurprising given the similarities of these instruments. Another explanation is that relative to the DAS28-ESR, the mRADAI-5 has the ability to capture foot disease activity as it includes questions that cover all joints as a whole [16]. Importantly, RADAI-F5 and mRADAI-5 scores were not perfectly correlated, and 72.5% agreement was observed between respective disease activity categories for each instrument, suggesting they capture local and global disease respectively.

The RADAI-F5 demonstrated very good reliability characteristics in terms of 1-week reproducibility of summary scores and agreement between baseline and 1-week disease categories. However, the SDC95 value (2.69) derived from the SEM (0.97) was larger than anticipated and exceeded the preliminary MID value (1.16) which was obtained via the distribution method. This means that if an individual patient has a change score as large as the preliminary MID, we cannot be 95% confident that this change is not due to measurement error [33]. This may be explained by the presence of outliers (Figure 3, and Figures 4-8 and Table 4 in supplementary material) rather than systematically large change scores. Inspection of outliers suggested a tendency for larger change scores in those with shorter disease durations. Moreover, the RADAI-F5 summary change scores appeared to be predominantly driven by larger change scores for individual items 4 (general foot health) and 5 (morning stiffness in the feet), which may be more unstable over shorter periods of time. This suggests that the instrument stability period of 1-week adopted here may be too long to evaluate instrument in those whose disease may be unstable. Moreover, it is possible that

changes in therapy such as administration of intra-muscular steroid could have affected 1-week scores. Anchor-based derivations of RADAI-F5 MID are now planned as these are recommended and should be assigned the most weight when estimating MID [38,42]. We recommend that population specific RADAI-F5 MID values are derived for future longitudinal studies seeking to measure changes in foot disease over time.

A strength of the RADAI-F5 is that it appears to demonstrate a consistent pattern of good responsiveness and theoretical consistency for longitudinal measurement, as has been reported for other global disease activity PROMs [16,43]. We evaluated responsiveness using data from a pragmatic trial comparing customised versus prefabricated foot orthoses plus routine medical care in early RA [25]. Whilst the responsiveness results over a 6-month period are promising, medical care was not standardised and so drugs and dosages varied between participants. Nevertheless, theoretically consistent associations for change scores between the RADAI-F5 and alternative instruments were observed, suggesting it could be used longitudinally to measure changes in foot disease activity in early RA.

Preliminary foot disease category thresholds are proposed here to enhance applicability in routine clinical care in line with conventional categories adopted in other global disease activity PROMs and composite indices [24]. Whilst there is a relatively broad spectrum of RADAI-F5 scores in the study sample, we observed a negatively skewed distribution indicative of predominant moderate-to-high foot disease activity within the sample. As a result, there were proportionally fewer participants allocated to 'remission' and 'mild' categories than the 'moderate' and 'high' categories. We also acknowledge that in the absence of a gold-standard outcome measure for quantifying foot disease activity, the mRADAI-5 (a global index of disease activity) was the best available reference score for establishing foot disease activity categories. Whilst threshold cut-offs appear to have good face validity, further evaluation utilising alternative approaches such as receiver operating characteristic curves may be appropriate. Our future work will seek to confirm preliminary foot disease activity category thresholds reported here with greater focus on those with more established disease who are in remission or low global disease activity states. We will also seek to determine whether treatment regimens would frequently be changed due to detection

of persistent moderate foot disease in patients who are in low disease activity or remission states according to DAS28-ESR.

We have confirmed the RADAI-F5 is valid, reliable, responsive, acceptable to patients, and potentially feasible for use in clinical practice in terms of ease of completion, quick scoring, and interpretability. The RADAI-F5 is freely available for use from www.gcu.ac.uk/centreforliving/radai-f5 and is recommended for use by rheumatologists and/or rheumatology nurse specialists alongside composite global disease activity indices; and Allied Health Professionals such as podiatrists and physiotherapists involved in delivering non-medical foot care in RA.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article and/or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. JW conceived the study and all authors contributed to the design of the study. GJH, DP, and AM were substantially involved in the collection of study data. GJH, AH, JW and MS were responsible for undertaking psychometric and statistical analyses.

REFERENCES

1. Wilson O, Hewlett S, Woodburn J, Pollock J, Kirwan J. Prevalence, impact and care of foot problems in people with rheumatoid arthritis: results from a United Kingdom based cross-sectional study. *J Foot Ankle Res* 2017;10:46.
2. van der Leeden M, Steultjens MPM, van Schaardenburg D, Dekker J. Forefoot disease activity in rheumatoid arthritis patients in remission: results of a cohort study. *Arthritis Res Ther* 2010;12:R3.
3. Michelson J, Easley M, Wigley FM, Hellmann D. Foot and ankle problems in rheumatoid arthritis. *Foot Ankle Int* 1994;15:608–13.
4. Otter SJ, Lucas K, Springett K, Moore A, Davies K, Cheek L, et al. Foot pain in rheumatoid arthritis prevalence, risk factors and management: an epidemiological study. *Clin Rheumatol* 2010;29:255–271.

5. van der Leeden M, Steultjens MPM, Terwee CB, Rosenbaum D, Turner DE, Woodburn J, et al. A systematic review of instruments measuring foot function, foot pain, and foot-related disability in patients with rheumatoid arthritis. *Arthritis Rheum* 2008;59:1257-1269.
6. Grondal L, Tengstrand B, Nordmark B, Wretenberg P, Stark A. The foot: still the most important reason for walking incapacity in rheumatoid arthritis. *Acta Orthop* 2008;79:257-261.
7. Aletaha D, Smolen JS. The Simplified Disease Activity Index (SDAI) and Clinical Disease Activity Index (CDAI) to monitor patients in standard clinical care. *Best Pract Res Clin Rheumatol* 2007;21:663–675.
8. van der Leeden M, Fiedler K, Jonkman A, Dahmen R, Roorda LD, van Schaardenburg D, et al. Factors predicting the outcome of customised foot orthoses in patients with rheumatoid arthritis: a prospective cohort study. *J Foot ankle Res* 2011;4:8.
9. Prevoo MLL, van Riel PLCM, van T’hof MA, van Rijswijk MH, van Leeuwen MA, Kuper HH, et al. Validity and reliability of joint indices. a longitudinal study in patients with recent onset rheumatoid arthritis. *Br J Rheumatol* 1993;32:589-594.
10. Aletaha D, Smolen J. The Simplified Disease Activity Index (SDAI) and the Clinical Disease Activity Index (CDAI): A review of their usefulness and validity in rheumatoid arthritis. *Clin Exp Rheumatol* 2005;23 (Suppl. 39):S100-S108.
11. Wechalekar MD, Lester S, Hill CL, Lee A, Rischmueller M, Smith MD, et al. Active foot synovitis in patients with rheumatoid arthritis: unstable remission status, radiographic progression, and worse functional outcomes in patients with foot synovitis in apparent remission. *Arth Care Res* 2016;68:1616-1623.
12. Wechalekar MD, Lester S, Proudman SM, Cleland LG, Whittle SL, Rischmueller M, et al. Active foot synovitis in patients with rheumatoid arthritis: applying clinical criteria for disease activity and remission may result in underestimation of foot joint involvement. *Arthritis Rheum* 2012;64:1316-1322.
13. Van Tuyl LHD, Britsemmer K, Wells GA, Smolen JS, Zhang B, Funovits J, et al. Remission in early rheumatoid arthritis defined by 28 joint counts: limited consequences of residual disease activity in the forefeet on outcome. *Ann Rheum Dis* 2012;71:33-37.
14. Bakker MF, Jacobs JWG, Kruize AA, van der Veen MJ, van Booma-Frankfort C, Vreugdenhil SA, Bijlsma JWJ, Lafeber FPJG, Welsing PMJ. Misclassification of disease

- activity when assessing individual patients with early rheumatoid arthritis using disease activity indices that do not include joint of feet. *Ann Rheum Dis* 2012;71;6;830-835.
15. Stucki G, Liang M, Stucki S, Bruhlmann P, Michel BA. A self-administered rheumatoid arthritis disease activity index (RADAI) for epidemiological research. Psychometric properties and correlation with parameters of disease activity. *Arthritis Rheum* 1995;38:795–798.
 16. Leeb BF, Handl PM, Maktari A, Nothnagl T, Rintelen B. Patient-centred rheumatoid arthritis disease activity assessment by a modified RADAI. *J Rheumatol* 2008;35:1294-1299.
 17. Helliwell P, Reay N, Gilworth G, Redmond A, Slade A, Tennant A, Woodburn J. Development of a foot impact scale for rheumatoid arthritis. *Arthritis Rheum* 2005;53:418-422.
 18. Budiman-Mak E, Conrad KJ, Roach KE. The Foot Function Index: a measure of foot pain and disability. *J Clin Epidemiol* 1991;44:561–570.
 19. Walsmley S, Ravey M, Graham A, The LS, Williams AE. Development of a patient-reported outcome measure for the foot affected by rheumatoid arthritis. *J Clin Epidemiol* 2012;65:413-422.
 20. Wakefield RJ, Freeston JE, O'Connor P, Reay N, Budgen A, Hensor EMA, et al. The optimal assessment of the rheumatoid arthritis hindfoot: a comparative study of clinical examination, ultrasound and high field MRI. *Ann Rheum Dis* 2008;67:1678–1682.
 21. Tan YK, Ostergaard M, Conaghan PG. Imaging tools in rheumatoid arthritis: ultrasound vs magnetic resonance imaging. *Rheumatology* 2012;51:vii36-vii42.
 22. Ohrndorf S, Boer AC, Boeters DM, ten Brinck RM, Burmester GR, Kortekaas MC, et al. Do musculoskeletal ultrasound and magnetic resonance imaging identify synovitis and tenosynovitis at the same joints and tendons? A comparative study in early inflammatory arthritis and clinically suspect arthralgia. *Arthritis Res Ther* 2019;21:59.
 23. Dale J, Purves D, McConnachie A, McInnes I, Porter D. *Arth Care Res* 2014;66:19-26.
 24. Rintelen B, Haindl PM, Sautner J, Leeb BA, Deutsh C, Leeb BF. The rheumatoid arthritis disease activity index-5 in daily use. Proposal for disease activity categories. *J Rheumatol* 2009;36:918-924.
 25. Gallagher KS, Godwin J, Hendry GJ, Steultjens M, Woodburn J. A protocol for a randomised controlled trial of prefabricated versus customised foot orthoses for people with

- rheumatoid arthritis: the FOCOS RA trial [Foot Orthoses – Customised v Off-the-Shelf in Rheumatoid Arthritis]. *J Foot ankle Res* 2018;11:24.
26. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-334.
27. Bot SDM, Terwee CB, van der Windt DAWM, Bouter LM, Dekker J, de Vet HCW. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann Rheum Dis* 2004;63:335-341.
28. Anderson J, Caplan L, Yazdany J, Robbins ML, Neogi T, Michaud K, Saag KG, O'Dell JR, Kazi S. Rheumatoid arthritis disease activity measures: American College of Rheumatology recommendations for use in clinical practice. *Arthritis Care & Research* 2012;64:640-647.
29. Muradin I, van der Heide HJL. The foot function index is more sensitive to change than the Leeds Foot Impact Scale for evaluating rheumatoid arthritis patients after forefoot or hindfoot reconstruction. *Int Orthop* 2016;40:745-749.
30. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. *Anesth Analg* 2018;126:1763-1768.
31. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *J Clin Epidemiol* 1990;43:551–558.
32. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
33. De Vet HCW, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006;59:1033-1039.
34. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;8:307-310.
35. Menz HB, Auhl M, Ristevski S, Frescos N, Munteanu SE. Comparison of the responsiveness of the Foot Health Status Questionnaire and the Manchester Foot Pain and Disability Index in older people. *Health Qual Life Outcomes* 2014;12:158.
36. Guyatt G, Walter S, Norman G: Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987;40:171–178.
37. Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. *BMJ Open* 2015;5:e007765.

38. Mokkink LB, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, de Vet HCW, Terwee CB. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs): user manual. February 2018: https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018.pdf.
39. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–1905.
40. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539-549.
41. Woodburn J, Hennessy K, Steultjens MPM, McInnes I, Turner DE. Looking through the 'window of opportunity': is there a new paradigm of podiatry care on the horizon in early rheumatoid arthritis? *J Foot Ankle Res* 2010;3:8.
42. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes* 2006;4:70.
43. Salaffi F, Ciapetti A, Gasparini S, Carotti M, Bombardieri S, on behalf of the “NEW INDICES” study group. The comparative responsiveness of the patient self-report questionnaires and composite disease indices for assessing rheumatoid arthritis activity in routine care. *Clin Experimental Rheumatol* 2012;30:912-921.

TABLES

Table 1.

Characteristics	Descriptive statistics
Participants, n	142
Sex, % female	72.5
Age in years, median (IQR)	55.5 (50-62) years
Disease duration in months, median (IQR)	10 (3.6-20.8) months
RADAI-F5 (0-10), median (IQR)	5.2 (3.2-7.3)
mRADAI-5 (0-10), median (IQR)	5.4 (3.4-7.2)
DAS28-ESR (0-10), median (IQR)	4.1 (2.84-4.92)
FFI (0-100), median (IQR)	45.9 (23.2-58.03)
FIS_IF (0-21), median (IQR)	13 (8.0-15.25)
FIS_AP (0-30), median (IQR)	16.0 (6.5-22.0)

Table 2.

Measure	Correlation coefficient (95%CI)	Strength of association	A priori hypothesis for association
Construct validity [†]			
mRADAI-5 (n=133)	0.79 (0.73-0.85)**	Strong	Moderate
DAS28-ESR (n=106)	0.38 (0.26-0.57)**	Weak	Weak
FFI (n=132)	0.71 (0.62-0.79)**	Strong	Moderate
FIS_IF (n=125)	0.69 (0.66-0.82)**	Moderate	Moderate
FIS_AP (n=124)	0.48 (0.37-0.63)**	Moderate	Moderate
Longitudinal validity [‡]			
mRADAI-5 (n=63)	0.66 (0.41-0.74)**	Moderate	Moderate
DAS28-ESR (n=60)	0.33 (0.04-0.52)*	Weak	Weak
FFI (n=64)	0.60 (0.47-0.77)**	Moderate	Moderate

FIS_IF (n=63)	0.43 (0.01-0.56)**	Moderate	Weak
FIS_AP (n=63)	0.19 (-0.01-0.49)	Weak	Weak
** p<0.01, * p<0.05, †using baseline scores, ‡using 0-6 month change scores			

Table 3.

	RADAI-F5 Disease Activity Categories			
	Remission (0-1), n=11	Mild (<1-3.6), n=15	Moderate (>3.6-5.7), n=42	High (>5.7-10), n=65
Female, %	75	62.1	89.2	33.3
Age, years	63 (47-74)	59 (53-69)	54.5 (49-59)	55 (50-61)
Disease duration, months	144 (48-246)	10 (3-19)	7 (4-14)	11 (3-22)
DAS remission, n (%)	3 (2.9)	10 (9.5)	6 (5.7)	3 (2.9)
DAS low disease, n (%)	0 (0)	2 (1.9)	1 (1.0)	6 (5.7)
DAS moderate disease, n (%)	0 (0)	10 (9.5)	14 (13.3)	26 (24.8)
DAS high disease, n (%)	0 (0)	1 (1.0)	8 (7.6)	15 (14.3)
DAS28-ESR	1.76 (0.8-1.8)	2.91 (2.3-4.4)	4.13 (3.0-5.1)	4.33 (3.6-5.3)
mRADAI-5	0.6 (0.2-2.9)	3.2 (2.3-5.2)	5.2 (3.8-6.4)	7.2 (6.0-7.9)
FFI	4.8 (0-10)	17.9 (9.6-43.7)	35.8 (29.4-47.9)	56.9 (48.3-69.2)

FIS_IF	0 (0-3.25)	8 (6-11.75)	12 (11-15)	15 (13-17)
FIS_AP	1.5 (0-8.25)	10 (4-15)	14 (10-22)	19 (16-24)

FIGURE LEGENDS

Figure 1. Scatterplots demonstrating convergent and divergent validity for RADAI-F5 associations with (A) mRADAI-5 (n=133) and (B) DAS-28 ESR scores (n=106).

Figure 2. Bland-Altman plot for the RADAI-F5 summary score, with data points illustrating average scores from 2 time points separated by 1 week (n=84) against the difference between the corresponding baseline and 1 week scores, bias (solid line representing the average SD for baseline to 1 week scores for all participants), bias upper and lower 95% CI limits (light grey wide dashed line), and upper and lower LOA (dark grey narrow dashed line).

Figure 3. Bar charts of participant responses (n=130) to RADAI-F5 relevance (A) and readability/understanding (B) Likert scales.

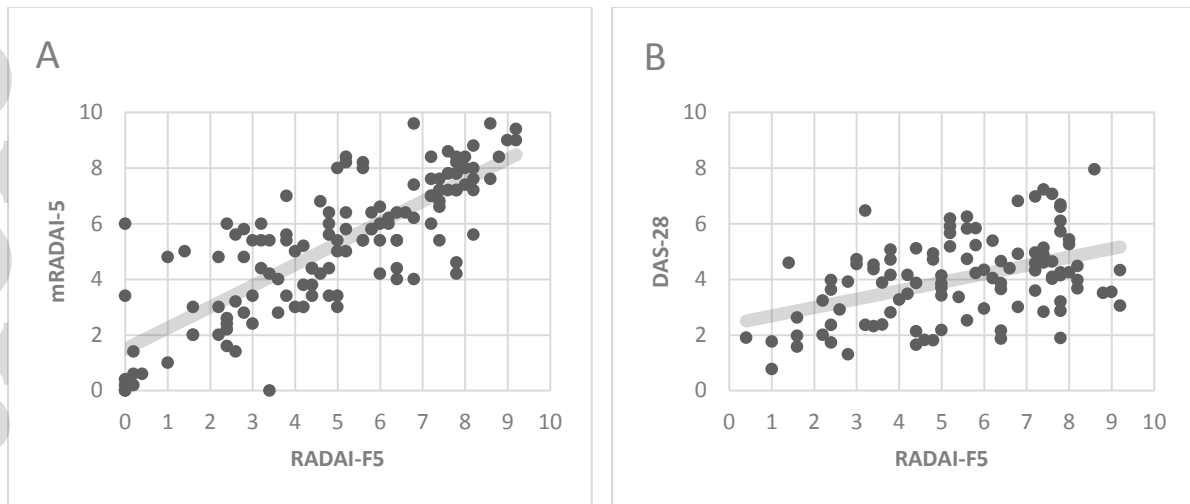


Figure 1.

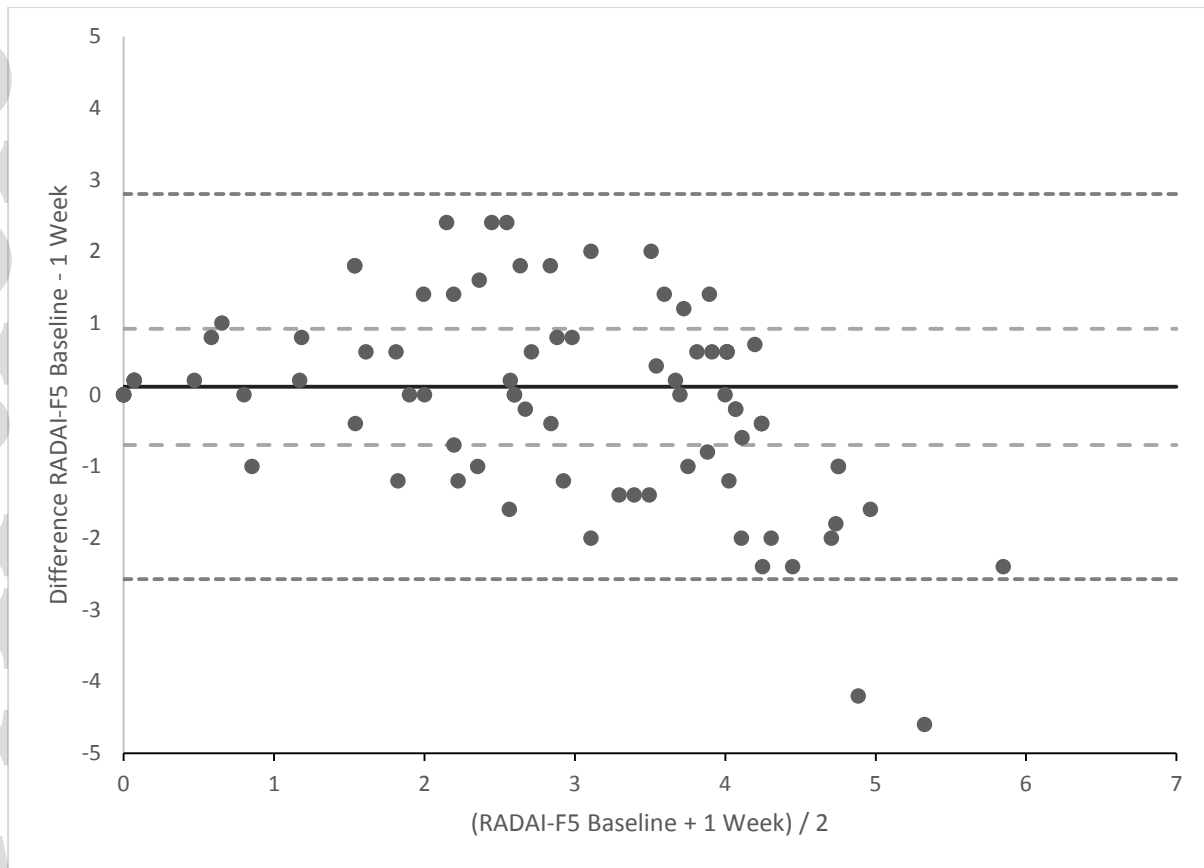


Figure 2.

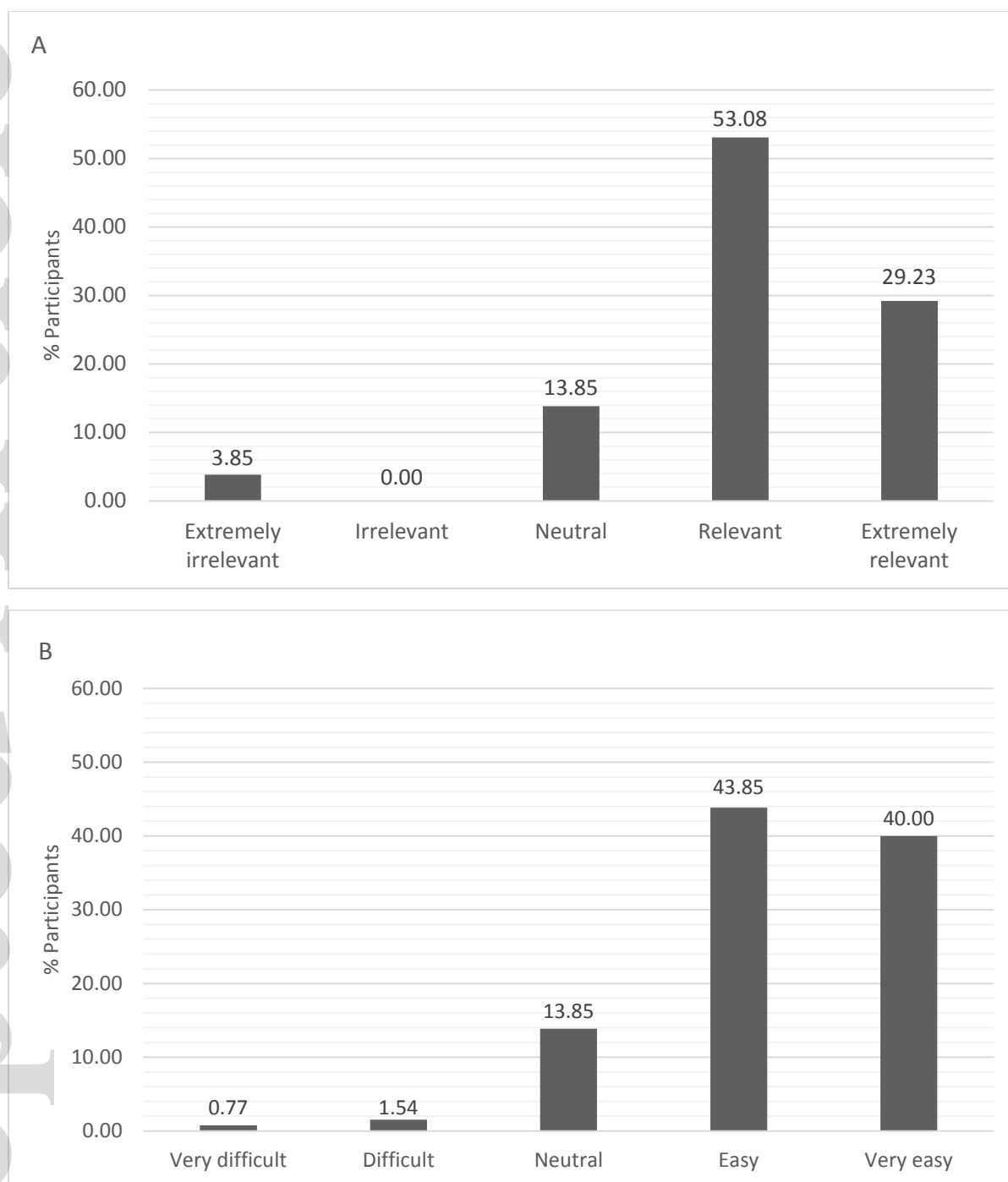


Figure 3.